Path–KL Friction: A Gauged KL–Projection Framework

Matthew Parsons
Independent Researcher
matthewgrantparsons@gmail.com

August 12, 2025

1 Introduction

Many engineered and natural systems generate trajectories (paths). We seek a principled measure of how far an *observed* system deviates from an idealized, *frictionless* baseline. Our construction produces (i) a *friction vector* $\mathbf{F} \in [0,1]^K$ from interpretable components (calibration, dispersion, latency, dynamic response, tail), (ii) a *scalar summary* via an entropic aggregator with tunable conservatism, and (iii) a *surrogate law* Q obtained via an I-projection on path space that matches declared constraints while minimizing KL divergence to a baseline R_0.

This paper synthesizes two strands: a measure-theoretic I-projection foundation [1–3] and a new axiomatic layer that pins *gauge*, invariances, stability, and reporting. We codify the rail (venue of action) and provide proofs of data-processing contraction and composition. The approach is domain-agnostic; a finance instantiation is provided for concreteness, and other domains follow by swapping semantics of path functionals.

2 Preliminaries and Objects

Let X be a measurable state space, horizon $T \in \mathbb{N}$, and $\Omega = X^{T+1}$ with canonical process $(X_t)_{t=0}^{T}$. We consider probability measures P, Q, R_0 on $(\Omega, \mathcal{F})$. P denotes the observed system; R_0 is the declared baseline; Q will denote an I-projection solution.

[Venue of action (*rail*)] A rail is a typed object $\mathcal{R} = (X, \mathcal{A}, \Pi, \Phi, \mathcal{E}, \mathfrak{T})$: state space X; action set $\mathcal{A}$; admissible policy class Π; stochastic transition kernel Φ; exogenous inputs $\mathcal{E}$; and a timing/latency operator $\mathfrak{T}$. Each instance induces a path law $P_{\mathcal{R},\pi}$ on Ω. We assume $P_{\mathcal{R},\pi} \ll R_0$ for all $\pi \in \Pi$.

[Friction vs. loss function vs. loss] *Friction* is a vector $\mathbf{F}_{\mathcal{R},\tau} \in [0,1]^K$ of normalized components measuring undesirable deviations over window τ. A *loss function* is the functional used in inference (here: KL divergence). A *loss* is a realized domain cost under a law Q (e.g., dollars, latency).

3 The Axiomatic and Gauge Framework

3.1 Gauge Definition

A central tenet of this framework is that no measurement is meaningful without a clear, declared context. We formalize this context as a *gauge* object. [Gauge $\mathcal{G}$] A *gauge* is the tuple

$$\mathcal{G} = \big(R_0,\ \mathsf{units},\ \mathsf{binning},\ B_\Delta,\ q,\ \tau_{\mathrm{norm}},\ h,\ \sigma\text{-}\mathsf{window},\ \mathbf{w},\ \theta,\ \alpha_0\big),$$

fixing the baseline, measurement units, discretisation rules, and all hyper-parameters. Any reported friction vector $\mathbf{F}$ or scalar F_Σ *must* cite $\mathcal{G}$.

3.2 Canonical Gauge Families

- ▷ **Baseline-anchored:** $\tau_{\mathrm{norm}} := \mathrm{CVaR}_q^{R_0}(\Delta^+)$; B_Δ via Freedman–Diaconis; h at the decorrelation time of R_0.
- ▷ **Budget-anchored:** τ_{norm} from a risk budget; B_Δ fixed by the reporting standard; h from SLOs.

3.3 Axioms of Path-KL Friction

The framework's behavior is governed by the following axioms. [Attainability & Regularity] The constraint set $\mathcal{L}_c$ is non-empty, the target c lies in the interior of the moment set, the statistics g are integrable, and the log-moment generating function $\psi(\lambda)$ is finite and smooth in a neighborhood of its maximizer.

[Boundedness & Monotonicity] Each component $F_k \in [0, 1]$ and is non-decreasing in its underlying undesirable statistic.

[Gauge Anchoring & Invariance] All scales are declared in $\mathcal{G}$; admissible affine re-parameterisations of the constraints leave the core friction unchanged.

[Stability] With the prescribed estimation policy (e.g., Dirichlet–shrinkage entropy, RU-CVaR), each estimator $\widehat{F}_k$ enjoys finite-sample deviation bounds.

[Data-Processing Contraction] Coarse-graining via any measurable map Φ cannot increase the minimum KL friction (see Appendix 12).

[Composition] For independent rails under aligned gauges, the minimal KL divergence is additive.

[Aggregation Coherence] The scalar summary $F_\Sigma(\theta) = \theta^{-1} \log \sum_k e^{\theta w_k F_k}$ is monotone, convex, and interpolates between a weighted mean and max.

4 Friction Components and Mapping to Constraints

We define interpretable components $F_k \in [0, 1]$ (finance semantics are used for illustration). Each component includes its g/c mapping to ensure linearity of constraints.

Predictive calibration. With calibrated scores $\hat{p}_i$ and outcomes y_i, define F_{predict} via a proper-scoring-rule-derived calibration error normalized to $[0,1]$. Constraint: match binwise calibration moments $g_{1,b} = \mathbb{1}\{\hat{p} \in \text{bin } b\}(\hat{p} - y)$ to $c_{1,b} \approx 0$.

Slippage dispersion (entropy). Let binned slippage $\tilde{\Delta} \in \{1, \ldots, B_\Delta\}$ with empirical histogram $\hat{\pi}(b)$. Define $F_{\text{disp}} = H(\tilde{\Delta})/\log B_\Delta$. Constraint: $g_{2,b}(\omega) = \mathbb{1}\{\tilde{\Delta}(\omega) = b\}$ with $c_{2,b} = \hat{\pi}(b)$ for all b; use Dirichlet shrinkage to stabilize estimates [4].

Latency. With delays $L_i \geq 0$ and scale B_L (from gauge), set $F_{\text{latency}} = \mathbb{E}[\, L/B_L \,] \wedge 1$. Constraint: first moment of the delay distribution.

Dynamic response. For post-action price move $\Delta P_{\text{post},i}$ at horizon h and pre-window volatility $\sigma_i > 0$, define $F_{\text{dyn}} = \frac{1}{N}\sum_{i=1}^{N} \min\{|\Delta P_{\text{post},i}|/\sigma_i,\, 1\}$. Constraint: match the mean of these capped ratios.

Tail risk. For loss-coded slippage Δ and quantile $q \in (0,1)$, define $F_{\text{tail}} = \min\{\text{CVaR}_q(\Delta^+)/\tau_{\text{norm}},\, 1\}$, with τ_{norm} from the gauge. Constraint: Rockafellar–Uryasev convex reformulation [5].

Aggregation. The scalar summary is $F_\Sigma(\theta) = 1/\theta \log\Big(\sum_{k=1}^{K} \exp\big(\theta\, w_k F_k\big)\Big)$, which is a certainty-equivalent under exponential utility [6].

5 I-Projection and Geometric Decomposition

Given path functionals $g : \Omega \to \mathbb{R}^K$ and target statistics $c \in \mathbb{R}^K$, we define the constraint set

$$\mathcal{L}_c = \{\, Q \ll R_0 : \mathbb{E}_Q[\, g \,] = c \,\} \tag{1}$$

and solve the I-projection problem:

$$\inf_{Q \in \mathcal{L}_c} \; \mathrm{D}_{\mathrm{KL}}(Q \,\|\, R_0)\,. \tag{2}$$

Under Axiom 3.3, the solution exists and is unique.

[I-projection and Pythagorean identity] The unique minimizer is

$$\frac{\mathrm{d}Q^*}{\mathrm{d}R_0}(\omega) = \exp\Big((\lambda^*)^\top g(\omega) - \psi(\lambda^*)\Big), \qquad \nabla\psi(\lambda^*) = c. \tag{3}$$

Moreover, for any $Q \in \mathcal{L}_c$,

$$\mathrm{D}_{\mathrm{KL}}(Q \,\|\, R_0) = \mathrm{D}_{\mathrm{KL}}(Q \,\|\, Q^*) + \mathrm{D}_{\mathrm{KL}}(Q^* \,\|\, R_0)\,. \tag{4}$$

See Appendices 10 and 11. Equation (3) identifies the solution as an exponential-family member with sufficient statistics g. Equation (4) geometrically separates deviation into a term measuring mismatch to the optimal family member Q and a term measuring the mismatch of the entire family to the baseline R_0 [2].

6 Key Properties and Consequences

6.1 Data-processing contraction

[Friction contracts under coarse-graining] For any measurable map $\Phi : \Omega \to \Omega'$, the resulting friction cannot increase. See Appendix 12.

6.2 Gauge invariance

[Admissible reparameterizations leave friction unchanged] If $g'(\omega) = Ag(\omega) + b$ with A invertible and $c' = Ac + b$, then $\mathcal{L}_{c'} = \mathcal{L}_c$ and the minimum KL is unchanged.

6.3 Composition laws

[Product rails: KL additivity] For two independent rails, the minimum KL divergence of the composite system is the sum of the individual minimum KL divergences. [Sketch] Follows from the properties of product exponential families and the additivity of their log-partition functions.

6.4 Stability highlights

Entropy. With Dirichlet shrinkage, Paninski [4] provides finite-sample deviation bounds for entropy estimates.

CVaR. For subexponential tails, the Rockafellar–Uryasev estimator has exponential-type concentration [7].

Dynamic response. For bounded ratios, Hoeffding's inequality [8] applies; for heavy tails, robust means are preferred [9].

7 Implementation and Reporting

Inference. The I-projection is solved via its dual problem, $\sup_\lambda\{\lambda^\top c - \psi(\lambda)\}$, typically with gradient-based methods.

Reporting standard. Publish the full vector $\mathbf{F}$ with 95% CIs; F_Σ at two temperatures; the exact gauge tuple $\mathcal{G}$; and a sensitivity table across baselines.

8 Related Connections

KL control / control as inference. There is a close relationship to linearly solvable control problems with KL penalties [10–12].

Schrödinger bridges. Entropic interpolation on path space connects to minimum-KL fitting under marginal constraints [13].

9 Conclusion

We have integrated a measure-theoretic KL projection program with a rigorous axiomatic and gauge specification, yielding a stable, comparable, and reproducible notion of system friction on path space. The framework scales across domains, provides invariances and composition rules, and supports policy-relevant reporting. Future work includes sharp equivalences to Schrödinger bridges for specific templates and end-to-end exemplars in non-finance domains.

9.1 Standing assumptions (A1–A4)

Let $(\Omega, \mathcal{F})$ be the path space with baseline law R_0. Let $g : \Omega \to \mathbb{R}^K$ be the vector of sufficient statistics (component functionals) and $c \in \mathbb{R}^K$ the target constraint vector. Define the constraint set and moment set

$$\mathcal{L}_c\{\, Q \ll R_0 :\ \mathbb{E}_Q[\, g\,] = c\,\}, \qquad \mathcal{M}\{\, \mathbb{E}_Q[\, g\,] : Q \ll R_0\,\}.$$

Let the log–moment generating function (log–Laplace / cumulant)

$$\psi(\lambda) \log \mathbb{E}_{R_0}\Big[\, \exp\big(\lambda^\top g(\omega)\big)\,\Big], \qquad \lambda \in \mathbb{R}^K.$$

(A1) Measurability & integrability. Each component of g is $\mathcal{F}$–measurable and $g \in L^1(R_0)$. There exists an open neighborhood $U \subset \mathbb{R}^K$ of 0 such that $\mathbb{E}_{R_0}\big[\exp(\lambda^\top g)\big] < \infty$ for all $\lambda \in U$ (i.e., ψ is finite near 0).

(A2) Attainability & interiority. The constraint set is nonempty: $\mathcal{L}_c \neq \varnothing$. Moreover the target lies in the (relative) interior of the moment set: $c \in \mathrm{ri}(\mathcal{M})$.

(A3) Regularity of the dual. ψ is finite and $\mathcal{C}^2$ on some open convex set $V \subset \mathbb{R}^K$ with $U \subseteq V$, and is strictly convex on V. The gradient map $\nabla\psi : V \to \mathrm{ri}(\mathcal{M})$ is onto.

(A4) Nondegeneracy / identifiability. At the optimizer λ^* (when it exists), the Hessian is positive definite: $\nabla^2\psi(\lambda^*) = \mathrm{Cov}_{Q^*}(g)$ is invertible. Equivalently, the components of g are not affinely dependent under Q^*.

Consequences used in T1/T2/T3. Under (A1)–(A4), the I–projection exists and is unique; the optimizer is exponential-family:

$$\frac{dQ^*}{dR_0}(\omega) = \exp\big(\lambda^{*\top} g(\omega) - \psi(\lambda^*)\big), \quad \nabla\psi(\lambda^*) = c,$$

and the Pythagorean identity holds for all $Q \in \mathcal{L}_c$:

$$\mathrm{D}_{\mathrm{KL}}(Q\|R_0) \ =\ \mathrm{D}_{\mathrm{KL}}(Q\|Q^*) \ +\ \mathrm{D}_{\mathrm{KL}}(Q^*\|R_0).$$

9.2 Componentwise KL decomposition

We apportion the total friction

$$D \;\mathrm{D}_{\mathrm{KL}}(Q^*\|R_0) \;=\; \lambda^{*\top} c \;-\; \psi(\lambda^*)$$

to the K declared components in a way that is gauge-coherent, order-free, and reproducible.

Canonical, order-free split. Turn all constraints on simultaneously along a straight line in constraint space:

$$c(t) \;\; t\,c, \qquad t \in [0,1].$$

For each t, solve the dual for $\lambda(t)$ via

$$\lambda(t) \in \arg\max_{\lambda}\{\lambda^\top c(t) - \psi(\lambda)\}, \qquad \nabla\psi(\lambda(t)) = c(t),$$

and define Q_t by the corresponding exponential tilt. Write

$$D(t) \;\; \mathrm{D}_{\mathrm{KL}}(Q_t\|R_0) \;=\; \lambda(t)^\top c(t) - \psi(\lambda(t)).$$

Differentiating and using $\nabla\psi(\lambda(t)) = c(t)$ gives

$$\frac{d}{dt}D(t) \;=\; \lambda(t)^\top \frac{dc(t)}{dt} \;=\; \lambda(t)^\top c \;=\; \sum_{k=1}^{K} \lambda_k(t)\, c_k.$$

Integrating from 0 to 1 yields

$$D \;=\; \int_0^1 \lambda(t)^\top c\, dt \;=\; \sum_{k=1}^{K} \Big(c_k \int_0^1 \lambda_k(t)\, dt\Big).$$

Definition. For each component k,

$$\boxed{\Delta D_k \;\; c_k \int_0^1 \lambda_k(t)\, dt, \qquad k = 1, \ldots, K.}$$

Properties. (i) *Efficiency:* $\sum_k \Delta D_k = D$. (ii) *Gauge coherence:* all terms are defined in the declared gauge G. (iii) *Order-free:* no component is privileged; the path $c(t) = tc$ treats them symmetrically. (iv) *Sign/monotonicity:* for undesirable statistics with $c_k \geq 0$, $\lambda_k(t)$ has the sign that increases D, so $\Delta D_k \geq 0$.

Computation. Choose quadrature nodes $0 = t_0 < \cdots < t_J = 1$ (e.g., Gauss–Legendre or Simpson). For each t_j, solve the dual for $\lambda(t_j)$ and set

$$\Delta D_k \;\approx\; c_k \sum_{j=0}^{J} w_j\, \lambda_k(t_j), \qquad D \;\approx\; \sum_{k=1}^{K} \Delta D_k,$$

with standard weights w_j. Under (A1)–(A4), $t \mapsto \lambda(t)$ is smooth; $J \in [8, 32]$ is sufficient in practice.

Alternatives (discrete splits). A nested (Pythagorean) split adds constraints one at a time along a chain $\varnothing = S_0 \subset S_1 \subset \cdots \subset S_K = \{1, \ldots, K\}$ and uses

$$D \;=\; \sum_{i=1}^{K} \big(\mathrm{D}_{\mathrm{KL}}(Q^{S_i} \| R_0) - \mathrm{D}_{\mathrm{KL}}(Q^{S_{i-1}} \| R_0)\big),$$

which is order-dependent. An order-invariant discrete split is the Shapley average over all permutations:

$$\Delta D_k^{\mathrm{sh}} \;=\; \sum_{S \subseteq [K] \setminus \{k\}} \frac{|S|!\,(K - |S| - 1)!}{K!} \Big(\mathrm{D}_{\mathrm{KL}}(Q^{S \cup \{k\}} \| R_0) - \mathrm{D}_{\mathrm{KL}}(Q^{S} \| R_0)\Big),$$

requiring up to 2^K I-projections (tractable for K=6) and satisfying efficiency, symmetry, and null-player axioms.

10 Proof of Theorem T1: Existence, Uniqueness, Exponential Form

Throughout let R_0, g, c, $\mathcal{L}_c$ (defined in (1)) and $\psi(\lambda) = ...mbda^\top g)]$ satisfy Assumptions (A1)–(A4) in the main text.

10.1 Lemma L1: Existence

There exists $Q^* \in \mathcal{L}_c$ attaining $\inf_{Q \in \mathcal{L}_c} \mathrm{D}_{\mathrm{KL}}(Q \,\|\, R_0)$.

Let $\{Q_n\} \subset \mathcal{L}_c$ be a minimising sequence with $...Q^*[g] = c$, so $Q^* \in \mathcal{L}_c$ and attains the infimum.

10.2 Lemma L2: Uniqueness

If $Q_1, Q_2 \in \mathcal{L}_c$ both attain $\inf_{Q \in \mathcal{L}_c} \mathrm{D}_{\mathrm{KL}}(Q \,\|\, R_0)$, then $Q_1 = Q_2$.

Strict convexity of $\mathrm{D}_{\mathrm{KL}}(\cdot \,\|\, R_0)$ in its first argument yields uniqueness: if $Q_\alpha = \alpha Q_1 + (1-\alpha) Q_2$ with $\alpha \in (0,1)$ also lies in $\mathcal{L}_c$, then $\mathrm{D}_{\mathrm{KL}}(Q_\alpha \,\|\, R_0) < \alpha \mathrm{D}_{\mathrm{KL}}(Q_1 \,\|\, R_0) + (1-\alpha)\mathrm{D}_{\mathrm{KL}}(Q_2 \,\|\, R_0)$, contradicting optimality unless $Q_1 = Q_2$.

10.3 Lemma L3: Exponential Form

The unique minimiser Q^* satisfies

$$\frac{\mathrm{d}Q^*}{\mathrm{d}R_0}(\omega) = \exp\big(\lambda^\top g(\omega) - \psi(\lambda)\big), \tag{5}$$

for some $\lambda \in \mathbb{R}^K$ with $\nabla\psi(\lambda) = c$.

Fenchel duality gives $\inf_{Q \in \mathcal{L}_c} \mathrm{D}_{\mathrm{KL}}(Q \,\|\, R_0) = \sup_{\lambda \in \mathbb{R}^K} \{\lambda^\top c - \psi(\lambda)\}$. Assumptions (A3–A4) ensure the supremum is attained at some λ^* with $\nabla\psi(\lambda^*) = c$. The exponential tilting $\frac{\mathrm{d}Q^*}{\mathrm{d}R_0} = \exp\big(\lambda^{*\top} g - \psi(\lambda^*)\big)$ realises the dual optimum and is therefore primal optimizer. Uniqueness follows from Lemma 10.2.

Conclusion of Theorem T1. Lemmas 10.1–10.3 establish existence, uniqueness, and the explicit exponential-family representation of the I–projection.

11 Proof of Theorem T2: Pythagorean Identity

Let Q^* be the unique I–projection from Appendix 10, satisfying

$$\frac{\mathrm{d}Q^*}{\mathrm{d}R_0}(\omega) = \exp\big(\lambda^{*\top} g(\omega) - \psi(\lambda^*)\big),$$

with $\nabla\psi(\lambda^*) = c$. For any $Q \in \mathcal{L}_c$ (defined in (1)) with $\mathrm{D}_{\mathrm{KL}}(Q \,\|\, R_0) < \infty$ we prove

$$\mathrm{D}_{\mathrm{KL}}(Q \,\|\, R_0) = \mathrm{D}_{\mathrm{KL}}(Q \,\|\, Q^*) + \mathrm{D}_{\mathrm{KL}}(Q^* \,\|\, R_0)\,.$$

Part I: Orthogonality of the residual. Let $\mathcal{F} = \{\exp(\lambda^\top g - \psi(\lambda)) : \lambda \in \mathbb{R}^K\}$ denote the exponential family relative to R_0. The tangent space at Q^* is spanned by the centered statistics $g - \mathbb{E}_{Q^*}[g]$. For $t \in \mathbb{R}^K$, define Q_t via

$$\frac{\mathrm{d}Q_t}{\mathrm{d}R_0} \propto \exp\big((\lambda^* + t)^\top g - \psi(\lambda^* + t)\big).$$

Then $\mathbb{E}_{Q_t}[g] = c$ for all t and $Q_0 = Q^*$. Differentiating at $t = 0$ yields $\mathbb{E}_{Q^*}[g] = c$ and $\nabla^2\psi(\lambda^*) = \mathrm{Cov}_{Q^*}(g)$, establishing the usual "orthogonality" of the residual to the constraint manifold at Q^*.

Part II: Bregman decomposition of KL. Write $f = \mathrm{d}Q/\mathrm{d}R_0$ and $f^* = \mathrm{d}Q^*/\mathrm{d}R_0$. Then

$$\begin{aligned}\mathrm{D}_{\mathrm{KL}}(Q \,\|\, R_0) &= \mathbb{E}_{R_0}[\, f \log f \,] \\ &= \mathbb{E}_{R_0}\Big[\, f \log \tfrac{f}{f^*} \,\Big] + \mathbb{E}_{R_0}[\, f \log f^* \,]\,.\end{aligned}$$

The first term equals $\mathrm{D}_{\mathrm{KL}}(Q \,\|\, Q^*)$. For the second term, use the exponential form of f^*:

$$\begin{aligned}\mathbb{E}_{R_0}[\, f \log f^* \,] &= \mathbb{E}_Q[\log f^*\,] \\ &= \mathbb{E}_Q\Big[\, \lambda^{*\top} g - \psi(\lambda^*) \,\Big] \\ &= \lambda^{*\top}\mathbb{E}_Q[\,g\,] - \psi(\lambda^*) \\ &= \lambda^{*\top} c - \psi(\lambda^*) \\ &= \mathbb{E}_{R_0}[\, f^* \log f^* \,] \\ &= \mathrm{D}_{\mathrm{KL}}(Q^* \,\|\, R_0)\,.\end{aligned}$$

Hence

$$\mathrm{D}_{\mathrm{KL}}(Q \,\|\, R_0) \;=\; \mathrm{D}_{\mathrm{KL}}(Q \,\|\, Q^*) \;+\; \mathrm{D}_{\mathrm{KL}}(Q^* \,\|\, R_0)\,.$$

Conclusion. Combining the two parts yields the Pythagorean identity: $\mathrm{D}_{\mathrm{KL}}(Q \,\|\, R_0) = \mathrm{D}_{\mathrm{KL}}(Q \,\|\, Q^*) + \mathrm{D}_{\mathrm{KL}}(Q^* \,\|\, R_0)$. If $Q \neq Q^*$ then $\mathrm{D}_{\mathrm{KL}}(Q \,\|\, Q^*) > 0$ by strict convexity, so $\mathrm{D}_{\mathrm{KL}}(Q \,\|\, R_0) > \mathrm{D}_{\mathrm{KL}}(Q^* \,\|\, R_0)$.

12 Proof of Theorem T3: Data-Processing Contraction

Let $\Phi : \Omega \to \Omega'$ be measurable. Define push-forward measures

$$R_0' := R_0 \circ \Phi^{-1}, \qquad Q' := Q \circ \Phi^{-1}.$$

Define

$$g'(\omega') := \mathbb{E}_{R_0}\big[\, g \mid \Phi = \omega' \,\big]$$

so that, by the tower property and the definition of Q',

$$\mathbb{E}_{Q'}\big[\, g' \,\big] \;=\; \mathbb{E}_{Q}\big[\, \mathbb{E}_{R_0}[\, g \mid \Phi \,]\,\big] \;=\; \mathbb{E}_{Q}[\, g \,].$$

The constraint families are $\mathcal{L}_c$ in (1) and

$$\mathcal{L}_c' := \{\, Q' \ll R_0' : \mathbb{E}_{Q'}\big[\, g' \,\big] = c \,\}.$$

Lemma 1. Constraint compatibility

For any $Q \in \mathcal{L}_c$, the push-forward Q' lies in $\mathcal{L}_c'$. If $Q \in \mathcal{L}_c$ then $\mathbb{E}_Q[\, g \,] = c$ by definition. By the push-forward relation and the definition of g', $\mathbb{E}_{Q'}[\, g' \,] = \mathbb{E}_Q[\, \mathbb{E}_{R_0}[\, g \mid \Phi \,]\,] = \mathbb{E}_Q[\, g \,] = c$, hence $Q' \in \mathcal{L}_c'$.

Lemma 2. KL data-processing inequality

For every $Q \ll R_0$,

$$\mathrm{D}_{\mathrm{KL}}\big(Q' \,\big\|\, R_0'\big) \;\le\; \mathrm{D}_{\mathrm{KL}}(Q \,\|\, R_0)$$

(Cover–Thomas, Thm. 2.9).

Main inequality

Pick any $Q \in \mathcal{L}_c$. Then $Q' \in \mathcal{L}_c'$ by Lemma 1 and $\mathrm{D}_{\mathrm{KL}}(Q' \,\|\, R_0') \le \mathrm{D}_{\mathrm{KL}}(Q \,\|\, R_0)$ by Lemma 2. Taking infima over the respective families gives

$$\inf_{Q' \in \mathcal{L}_c'} \mathrm{D}_{\mathrm{KL}}\big(Q' \,\big\|\, R_0'\big) \;\le\; \inf_{Q \in \mathcal{L}_c} \mathrm{D}_{\mathrm{KL}}(Q \,\|\, R_0)\,.$$

A Case Study: SWIFT Cross-Border Payment Friction (UK→US vs. Asia→Africa)

Status and data provenance. This appendix accompanies the main paper. Quantitative inputs are taken from a single-source datasheet (v1; log-normal fits to median and $q_{0.99}$) consistent with CPMI SWIFT-gpi analytics and SWIFT summaries [14, 15]. Operating-hour effects follow CPMI guidance [16]. KPI targets are from the FSB progress report [17]. A maturity benchmark is from SWIFT [18].

A.1 Gauge G and Reporting Schema

Baselines (for sensitivity): E5 = Exp(12 h^{-1}) (mean 5 minutes), E15 = Exp(4 h^{-1}) (mean 15 minutes), LN_0 = thin-tailed log-normal with 5-minute mean.
Units: hours. **Scales:** B_L=24, B_σ=3, τ_{norm}=24, B_{idle}=12, $H_{\max}$=4.
Nesting order: cal → route → lat → disp → tail → dyn.
Aggregator: $F_\Sigma(\theta) = \frac{1}{\theta}\log\left(\frac{1}{K}\sum_{k=1}^{K} e^{\theta F_k}\right)$ with K=6 and $\theta \in \{1, 5\}$.
Publishing rule: Report G; the bounded friction vector $F \in [0,1]^6$; $F_\Sigma(\theta)$ at $\theta \in \{1,5\}$; the $D_{KL}(\cdot\|\cdot)$ decomposition by component; multipliers/fit parameters; and baseline sensitivity. Totals are gauge-relative.

A.2 Quantile Fits (Corridor Distributions)

Each corridor $X \geq 0$ is modeled as $LN(\mu, \sigma)$ fitted to the observed median and $q_{0.99}$ consistent with CPMI SWIFT-gpi medians and slow-corridor tails reported by SWIFT [14, 15].

Table 1: Corridor fits (median and $q_{0.99}$ → LN parameters).

Corridor	**Median (h)**	$q_{0.99}$ **(h)**	μ	σ	**Mean (h)**	**Std (h)**
UK→US	0.20	3	−1.609438	1.164078	0.393805	0.667968
Asia→Africa	22	96	3.091042	0.633313	26.885442	18.885837

A.3 Bounded Component Scores ($F \in [0,1]^6$)

Definitions: $F_{\text{lat}} = \min\{E[X]/24, 1\}$; $F_{\text{disp}} = \min\{\text{Std}[X]/3, 1\}$; $F_{\text{tail}} = \min\{E[(X-24)_+]/24, 1\}$; $F_{\text{dyn}} = \min\{\text{idle}/12, 1\}$; $F_{\text{route}} = \min\{(\mathbb{E}[\text{hops}]-1)/(H_{\max}-1), 1\}$. Dynamic idle reflects non-overlapping hours and batch windows [16]. F_{cal} is a placeholder calibration score for prior-predictive alignment.

Table 2: Bounded component scores $F \in [0,1]^6$.

Corridor	F_{cal}	F_{route}	F_{lat}	F_{disp}	F_{tail}	F_{dyn}
UK→US	0.0500	0.0000	0.01641	0.22266	0.00001	0.01667
Asia→Africa	0.1500	0.1000	1.00000*	1.00000*	0.32762	0.66667

Defs: $F_{\text{lat}} = \min\{E[X]/24, 1\}$, $F_{\text{disp}} = \min\{\text{Std}[X]/3, 1\}$, $F_{\text{tail}} = \min\{E[(X-24)_+]/24, 1\}$, $F_{\text{dyn}} = \min\{\text{idle}/12, 1\}$. * capped at 1 by G.

Aggregates:

$$F_\Sigma(1) = \begin{cases} 0.0542 & \text{(UK→US)} \\ 0.6097 & \text{(Asia→Africa)} \end{cases} \qquad F_\Sigma(5) = \begin{cases} 0.0697 & \text{(UK→US)} \\ 0.8037 & \text{(Asia→Africa)} \end{cases}$$

A.4 KL Decomposition (Baseline E5)

Total friction is $D = \mathrm{D_{KL}}(Q^* \,\|\, R_0)$ with $R_0 = \mathrm{Exp}(12\ \mathrm{h}^{-1})$. Component increments ($\Delta D_k$) are computed using the order-free path-integral method described in §9.2. Units: nats.

Table 3: Component ΔKL (nats), baseline $R_0 = \mathrm{Exp}(12\ \mathrm{h}^{-1})$.

Corridor	ΔD_{cal}	$\Delta D_{\mathrm{route}}$	ΔD_{lat}	ΔD_{disp}	ΔD_{tail}	ΔD_{dyn}	**Total**
UK→US	0.02769	0.00000	2.17265	0.06975	0.00000	0.00923	**2.27932**
Asia→Africa	0.00856	0.00571	315.84882	0.16736	0.01870	0.03806	**316.08721**

Central finding. Latency dominates. Asia→Africa: 315.84882/316.08721 = 99.9% of total friction is ΔD_{lat}. UK→US: 2.17265/2.27932 ≈ 95% is ΔD_{lat}. Dynamic friction in Asia→Africa has a large bounded score (F_{dyn}=0.667) but contributes only 0.038 nats to $\mathrm{D_{KL}}$—a rounding error next to latency.

Fit & multipliers. Tilt on X: $\lambda_X = 12 - 1/m$. UK→US: $\lambda_X = 9.4607\ \mathrm{h}^{-1}$; Gamma step $k = 0.3476$, $\theta = 1.1330$. Asia→Africa: $\lambda_X = 11.9628\ \mathrm{h}^{-1}$; Gamma step $k = 2.0266$, $\theta = 13.2665$.

A.5 Baseline Sensitivity (Totals Only)

Table 4: Baseline sensitivity of totals (nats).

Corridor	**E5** (5-min Exp)	**E15** (15-min Exp)	$\mathbf{LN}_0$ (mean 5-min)
UK→US	2.279	0.227	2.001
Asia→Africa	316.087	102.102	41.179

Interpretation: The absolute magnitude of $\mathrm{D_{KL}}(\cdot\|\cdot)$ is gauge-relative. Relaxing the baseline (E5 → E15 or LN_0) compresses totals—sharply for slow corridors—without changing the qualitative dominance of latency.

A.6 Implications

Primary recommendation: **attack baseline latency**. In Asia→Africa, almost all friction is the mean delay shift away from the 5-minute baseline. Routing simplification, extended operating hours, and process/queue redesign are necessary but *insufficient* unless they materially reduce end-to-end latency [16]. Against policy targets (75% within 1h; remainder within 1 day), current performance is far off in many corridors [17]. For contrast, see [18].

References

[1] Imre Csiszár. I-divergence geometry of probability distributions and minimization problems. *The Annals of Probability*, 3(1):146–158, 1975.

[2] Shun-ichi Amari and Hiroshi Nagaoka. *Methods of Information Geometry*, volume 191. American Mathematical Society, 2000.

[3] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley & Sons, 2nd edition, 2006.

[4] Liam Paninski. Estimation of entropy and mutual information. *Neural Computation*, 15(6): 1191–1253, 2003.

[5] R. Tyrrell Rockafellar and Stanislav Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2(3):21–41, 2000.

[6] Hans Föllmer and Alexander Schied. Convex measures of risk and trading constraints. *Finance and Stochastics*, 6(4):429–447, 2002.

[7] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, 2013.

[8] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[9] Olivier Catoni. Challenging the empirical mean and empirical variance: A deviation study. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 48(4):1148–1185, 2012.

[10] H. J. Kappen. Linear theory for control of nonlinear stochastic systems. *Physical Review Letters*, 95(20):200201, 2005.

[11] Emanuel Todorov. Efficient computation of optimal actions. *Proceedings of the National Academy of Sciences*, 106(28):11478–11483, 2009.

[12] Sergey Levine. Reinforcement learning and control as probabilistic inference: Tutorial and review. *arXiv preprint arXiv:1805.00909*, 2018.

[13] Christian Léonard. A survey of the schrödinger problem and some of its connections with optimal transport. *Discrete & Continuous Dynamical Systems-A*, 34(4):1533–1574, 2014.

[14] Thomas Nilsson, Regis Bouther, Michiel Van Acoleyen, and Lior Cohen. Swift gpi data indicate drivers of fast cross-border payments. Cpmi analytical note, Bank for International Settlements, Committee on Payments and Market Infrastructures (CPMI), February 2022. URL https://www.bis.org/cpmi/publ/swift_gpi.pdf. Based on about 20M SWIFT gpi transactions (Sep–Oct 2020); regional medians include about 12 minutes to North America and over 22 hours to North Africa.

[15] SWIFT. Is technology the panacea for today's cross-border payment frictions?, October 2023. URL https://www.swift.com/news-events/news/technology-panacea-todays-cross-border-payment-frictions. Explainer citing regional medians and beneficiary-leg delays; slowest routes over 2 days.

[16] Committee on Payments and Market Infrastructures. Extending and aligning payment system operating hours for cross-border payments. Cpmi papers d203, Bank for International Settlements, May 2022. URL https://www.bis.org/cpmi/publ/d203.pdf. Final report (12 May 2022): RTGS overlap, global settlement window; survey of 82 jurisdictions.

[17] Financial Stability Board. Annual progress report on meeting the targets for cross-border payments: 2024 report on key performance indicators. Technical report, Financial Stability Board, October 2024. URL https://www.fsb.org/uploads/P211024-3.pdf. Includes speed KPIs and 2027 targets (75% within 1 hour; remainder within 1 business day).

[18] SWIFT. Spotlight on speed: Where to focus for faster international payments. Technical report, SWIFT, October 2024. URL https://www.swift.com/sites/default/files/files/swift-spotlight-on-speed-report-october2024.pdf. Reports about 90% of cross-border payments reach the beneficiary bank within one hour (Oct 2024).